



19 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENT- UND
MARKENAMT

12 Patentschrift
10 DE 198 49 231 C 2

51 Int. Cl. 7:
G 01 N 24/08

21 Aktenzeichen: 198 49 231.6-33
22 Anmeldetag: 26. 10. 1998
43 Offenlegungstag: 18. 5. 2000
45 Veröffentlichungstag
der Patenterteilung: 31. 8. 2000

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

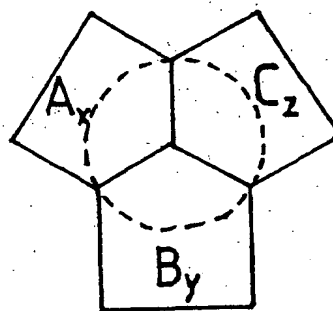
73 Patentinhaber:
Bruker Analytik GmbH, 76287 Rheinstetten, DE
74 Vertreter:
Kohler Schmid + Partner, 70565 Stuttgart

72 Erfinder:
Schröder, Harald, Dipl.-Chem.Dr., 79576 Weil am Rhein, DE; Neidig, Klaus-Peter, Dr., 76275 Ettlingen, DE

56 Für die Beurteilung der Patentfähigkeit in Betracht gezogene Druckschriften:
US-Z.: "J.Magn.Reson.", 89, (1990), S. 543-552;
US-Z.: "Analytical Chemistry News & Features", 1. Juni 1998, S. 378A, 385A u. 386A;
DE-Buch: "Ein- und Zweidimensionale NMR-Spektroskopie", H. Friebolin, 2. Aufl., 1992, VCH-Verlagsgesellschaft, Weinheim, S. 111-113 u. S. 244-249;
US-Z.: "J. Magn. Reson.", B 108, (1995), S. 94-98;

54 Verfahren zum Verifizieren der Synthese organischer Moleküle in der kombinatorischen Chemie

57 Verfahren zum Verifizieren der Synthese von aus Substrukturen aufgebauten organischen Molekülen der kombinatorischen Chemie, wobei die Substrukturen zumindest Teilbereiche der Ausgangssubstanzen enthalten, mit folgenden Schritten:
(A) Festlegung der Substrukturen der organischen Moleküle aufgrund der Ausgangssubstanzen und Zuordnen von Substrukturcodes (A_x , B_y , C_z , ...);
(B) Zuordnung eines Gesamtcodes für die jeweiligen Kombinationen von aus unterschiedlichen Substrukturen aufgebauten synthetisierten organischen Molekülen;
(C) Aufnahme von mehrdimensionalen NMR-Spektren der durch die jeweilige Kombination erhaltenen synthetisierten organischen Moleküle für zumindest eine Teilmenge der kombinatorisch möglichen Moleküle;
(D) eindeutige Zuordnung von Bereichen im NMR-Spektrum, in denen charakteristische Signale einer Substruktur vorkommen zu dieser Substruktur;
(E) Überprüfung der NMR-Spektren der Kombinationen auf das Vorhandensein der charakteristischen Signale in allen zugeordneten Bereichen;
(F) falls die Überprüfung für eine bestimmte Kombination ergibt, daß genau in den zugeordneten Bereichen die Signale der in ihrem Gesamtcodes enthaltenen Substrukturen im NMR-Spektrum beobachtet wurden, wird diese Kombination mit "wahr" charakterisiert.



DE 198 49 231 C 2

DE 198 49 231 C 2

Beschreibung

Die Erfindung betrifft ein Verfahren zum Verifizieren der Synthese von aus Substrukturen aufgebauten organischen Molekülen der kombinatorischen Chemie, wobei die Substrukturen zumindest Teilbereiche der Ausgangssubstanzen enthalten.

Verfahren der kombinatorischen Chemie zur automatischen Synthetisierung einer großen Zahl von neuen organischen Molekülen aus an sich bekannten Substrukturen und anschließende Verifikation des Erfolgs oder Mißerfolgs der Synthese sind beispielsweise bekannt aus dem Artikel "COMBINATORIAL" von A. W. Czarnik in der Zeitschrift Analytical Chemistry News & Features, Seiten 378 A bis 386 A, 1. Juni 1998.

Die kombinatorische Chemie ist ein Untergebiet der Chemie, das darauf abzielt, eine kleine Anzahl von chemischen Reagentien in allen durch ein gegebenes Reaktionsschema definierten Kombinationen zu synthetisieren, um eine große Anzahl von wohldefinierten Produkten zu erhalten.

Insbesondere besteht in der kombinatorischen Chemie ein großer Bedarf für NMR-Messungen mit hohem Durchsatz zur Verifizierung der synthetisierten Produkte. Die Auswertung der gemessenen NMR-Spektren wurde bisher "von Hand" durch hochspezialisierte Chemiker im wesentlichen intuitiv durchgeführt oder basierte auch auf im Ergebnis relativ ungenauen Modellrechnungen.

Aufgabe der vorliegenden Erfindung ist es daher, ein Verfahren der eingangs genannten Art vorzustellen, mit dem eine schnelle, reproduzierbare und zuverlässige Verifizierung einer großen Anzahl von synthetisierten Produkten der kombinatorischen Chemie ermöglicht wird.

Erfindungsgemäß wird diese Aufgabe durch folgende Verfahrensschritte gelöst:

- (A) Festlegung der Substrukturen der organischen Moleküle aufgrund der Ausgangssubstanzen und Zuordnen von Substrukturcodes (A_x, B_y, C_z, \dots);
- (B) Zuordnung eines Gesamtcodes für die jeweiligen Kombinationen von aus unterschiedlichen Substrukturen aufgebauten synthetisierten organischen Molekülen;
- (C) Aufnahme von mehrdimensionalen NMR-Spektren der durch die jeweilige Kombination erhaltenen synthetisierten organischen Moleküle für zumindest eine Teilmenge der kombinatorisch möglichen Moleküle;
- (D) eindeutige Zuordnung von Bereichen im NMR-Spektrum, in denen charakteristische Signale einer Substruktur vorkommen zu dieser Substruktur;
- (E) Überprüfung der NMR-Spektren der Kombinationen auf das Vorhandensein der charakteristischen Signale in allen zugeordneten Bereichen;
- (F) falls die Überprüfung für eine bestimmte Kombination ergibt, daß genau in den zugeordneten Bereichen die Signale der in ihrem Gesamtcodes enthaltenen Substrukturen im NMR-Spektrum beobachtet wurden, wird diese Kombination mit "wahr" charakterisiert.

Die zeitliche Abfolge der Verfahrensschritte (A) bis (F) läuft bevorzugt in der obengenannten Reihenfolge ab, kann jedoch auch in einer sinnvoll geänderten anderen Reihenfolge ablaufen. Denkbar ist beispielsweise, daß die Zuordnung von Signalbereichen der NMR-Spektren zu den einzelnen Substrukturen gemäß Schritt (D) aufgrund von Vorinformationen beispielsweise schon vor dem Schritt (A) erfolgt. Andere Verfahrensschritte wiederum, die auf vorhergehenden Verfahrensschritten zwingend aufbauen, können natürlicherweise nur in einer bestimmten Abfolge abgearbeitet werden. Beispielsweise kann die Überprüfung der NMR-Spektren in Schritt (E) erst dann erfolgen, wenn vorher die NMR-Spektren in Schritt (C) aufgenommen wurden.

Bei einer bevorzugten Variante des erfindungsgemäßen Verfahrens schließt sich an Schritt (F) folgender Schritt (G) an:

- (G) Falls die Überprüfung für eine bestimmte Kombination ergibt, daß in mindestens einem zugeordneten Bereich die charakteristischen Signale der in ihrem Gesamtcodes enthaltenen Substrukturen im NMR-Spektrum nicht beobachtet wurden, wird diese Kombination mit "falsch" charakterisiert.

Dadurch kann zunächst neben den positiv erkannten, wunschgemäß gelaufenen Synthesen, die in Schritt (F) mit "wahr" charakterisiert wurden, sehr schnell aus den übrigen Spektren diejenige Probe aussortiert werden, bei denen die Synthese nicht oder nicht vollständig funktioniert hat und daher zumindest eine der Signalgruppen im NMR-Spektrum nicht vorhanden ist. Für die als "falsch" charakterisierten Kombinationen kann anschließend entweder nochmals der Versuch einer Synthese aus den entsprechenden chemischen Ausgangssubstanzen gemacht werden, um sicherzustellen, daß die offenbar unvollständige Synthese beim ersten Mal nicht etwa auf einem Fehler im Syntheseverfahren beruht hat. Alternativ kann aber auch bei einer sehr großen Anzahl von zu bearbeitenden Kombinationen auf eine weitere Betrachtung der als "falsch" charakterisierten Kombinationen, die offenbar nicht ohne weiteres zu synthetisieren sind, ganz verzichtet werden.

In einer Weiterbildung kann sich an den Schritt (G) ein Schritt (H) anschließen:

- (H) Falls die Überprüfung für eine bestimmte Kombination ergibt, daß weder (F) noch (G) zutrifft, wird diese Kombination mit "unklar" charakterisiert.

Bei dieser Gruppe von Kombinationen waren offenbar alle Substrukturen an dem synthetisierten organischen Gesamtmolekül nachgewiesen worden, jedoch darüberhinaus auch noch weitere, die eigentlich nicht erwartet wurden. Hier könnte beispielsweise ein Hinweis auf unrichtig abgelaufene Syntheseverfahren oder anderweitige Meßfehler vorliegen, die beispielsweise zu Artefakten in den NMR-Spektren geführt haben könnten. Diese Fragen müßten bei Interesse anschließend von einem Spezialisten anhand der "unklaren" NMR-Spektren geklärt werden.

Ganz besonders bevorzugt ist eine Variante des erfindungsgemäßen Verfahrens, bei der die Schritte (D) und (E) durch

folgende Schritte realisiert werden:

- (D₁') Auswahl eines repräsentativen Satzes von Kombinationen, der alle interessierenden vorkommenden Substrukturen zumindest einmal enthält;
- (D₂') eindeutige Zuordnung von Bereichen mit charakteristischen Signalen der NMR-Spektren des repräsentativen Satzes zu den einzelnen Substrukturen;
- (D₃') falls sich nicht jeder Substruktur eindeutig ein Bereich von Signalen zuordnen läßt, erfolgt eine Auswahl eines geänderten Satzes von Kombinationen und Wiederholung ab Schritt (D₂');
- (E') Überprüfung der NMR-Spektren der restlichen, nicht dem repräsentativen Satz angehörigen Kombinationen auf das Vorhandensein aller zugeordneten Bereiche von Signalen.

Der repräsentative Satz von Kombinationen wird in der Regel erheblich kleiner sein als die Gesamtanzahl aller möglichen Kombinationen, so daß die Überprüfung der NMR-Spektren der restlichen Kombinationen erheblich beschleunigt werden kann. Voraussetzung hierfür ist natürlich, daß auch tatsächlich eine eindeutige Zuordnung jeder im Produktmolekül einer Kombination enthaltenen Substruktur zu einem Bereich von charakteristischen Signalen des repräsentativen Satzes möglich ist. Falls dies nicht der Fall sein sollte, muß in Schritt (D₃') der repräsentative Satz geändert werden und ein erneuter Versuch einer eindeutigen Zuordnung in Schritt (D₂') erfolgen. Erst wenn die Zuordnung eindeutig ist, können in Schritt (E') sämtliche restlichen NMR-Spektren mit Hilfe der zugeordneten Bereiche von Signalen des repräsentativen Satzes mit dem Ziel der Verifizierung der jeweiligen Synthese überprüft werden.

Eine Weiterbildung dieser bevorzugten Verfahrensvariante zeichnet sich dadurch aus, daß der repräsentative Satz in Schritt (D₁') durch Modifikation eines repräsentativen Satzes von Kombinationen aus einer früheren Meßreihe abgeleitet wird. Da in den meisten Fällen bereits Erfahrungen mit ähnlichen oder sogar gleichen Substrukturen vorliegen, können diese für einen "initial guess" für einen repräsentativen Satz von Kombinationen herangezogen werden, der mit höherer Wahrscheinlichkeit dem Kriterium einer eindeutigen Zuordnung von Signalbereichen zu den einzelnen Substrukturen genügen wird.

Vorzugsweise wird die Anzahl der im repräsentativen Satz enthaltenen Kombinationen minimiert, um die Überprüfung der restlichen NMR-Spektren nach dem erfindungsgemäßen Verfahren beschleunigen zu können. Falls die Substrukturen in Gruppen mit gleichem oder ähnlichem chemischen Verhalten aufgeteilt werden können, entspricht die kleinste denkbare Anzahl der im repräsentativen Satz enthaltenen Kombinationen der größten Anzahl von Substrukturen in einer Gruppe.

Die Zuordnung von Signalbereichen des repräsentativen Satzes in den NMR-Spektren zu den einzelnen Substrukturen in Schritt (D₂') kann bei Varianten des erfindungsgemäßen Verfahrens manuell erfolgen, wobei dann keine spezielle Auswertesoftware erforderlich wird.

Alternativ kann die Zuordnung aber auch automatisiert mittels eines Rechners erfolgen, der im Gegensatz zur "Handauswertung" erheblich schneller arbeitet, wobei teures, hochqualifiziertes Bedienungspersonal für die eigentliche Auswertung eingespart wird und die Überprüfung objektiv, mit sehr hoher Wahrscheinlichkeit fehlerfrei und reproduzierbar durch die Maschine erfolgt.

Bei einer weiteren bevorzugten Variante des erfindungsgemäßen Verfahrens erfolgt die Zuordnung von Signalbereichen zu den Substrukturen in Schritt (D) aufgrund von NMR-Spektren aus einer früheren Meßreihe und/oder aufgrund theoretisch berechneter Spektraldaten. Mit den vorhandenen Vorinformationen kann der oben beschriebene Auswertungsweg über einen repräsentativen Satz von Kombinationen abgekürzt werden und die Zuordnung von Signalbereichen zu den Substrukturen direkt erfolgen.

Vorzugsweise werden die zu synthetisierenden organischen Moleküle niedermolekular sein und in der Regel in einem Molekulargewichtsbereich von etwa 100 u bis etwa 2000 u liegen. Dies ist ein für die kombinatorische Chemie bevorzugter Massenbereich. Für die NMR-Messungen sind in diesem Bereich interessierende Substrukturen als Ausgangssubstanzen in ausreichenden Mengen vorhanden. Außerdem sind zweidimensionale NMR-Spektren in diesem Molekulargewichtsbereich noch relativ übersichtlich.

Besonders bevorzugt enthält das synthetisierte organische Gesamtmolekül einen im folgenden mit "Core" bezeichneten Abschnitt, der bei allen Kombinationen identisch vorhanden ist und folglich in den NMR-Spektren aufgrund der gleichen zugeordneten Bereiche von Signalen charakterisierbar ist, so daß es als interne Referenz im Sinne einer Standardintensität herangezogen werden kann. Das Core kann entweder vor der Synthese als eigenes Reagens zugegeben werden, wobei sich dann die anderen Substrukturen daran anlagern, oder es kann in einem Bindungsbereich der Substrukturen von diesen selbst als gemeinsamer Abschnitt des Moleküls gebildet werden.

Vorzugsweise ist das Core eine Substruktur, die zwischen zwei und sechs chemische Bindungsstellen aufweist. Damit bleibt die mögliche Kombinatorik hinreichend übersichtlich. Außerdem ist für diese Art von Substrukturen eine Vielzahl von handelsüblichen Substanzen als Core einsetzbar.

Bei einer ganz besonders bevorzugten Weiterbildung der oben beschriebenen erfindungsgemäßen Verfahrensvarianten sind die Substrukturen zu Gruppen zusammengefaßt, die jeweils einer Bindungsstelle des Core zugeordnet sind.

Die Anzahl der interessierenden Substrukturen sollte nicht wesentlich über 500 liegen, damit die Kombinatorik zuverlässig mit relativ hoher "Trefferwahrscheinlichkeit" ablaufen kann, andererseits aber deutlich über drei liegen, damit ein kombinatorischer Ansatz überhaupt sinnvoll wird.

Besonders bevorzugt werden die NMR-Spektren zweidimensional sein, insbesondere HSQC-Spektren (= hetero nuclear single quantum coherence, siehe z. B. J. Magn. Reson. B108, S.94-98 (1995)). Zweidimensionale NMR-Spektren lassen sich in endlich kurzen Meßzeiten in der Größenordnung von Minuten mit hinreichend genauer Auflösung im Gegensatz zu lediglich eindimensionalen Spektren erzeugen.

Ganz besonders bevorzugt sind auch Varianten des erfindungsgemäßen Verfahrens, bei denen das mehrdimensionale NMR-Spektrum die Atomkerne ¹H und ¹³C umfaßt. Damit sind die beiden wichtigsten Atomsorten der organischen Chemie, die nämlich dieses Untergebiet überhaupt definieren, enthalten.

Besonders einfach kann die Zuordnung von Signalbereichen der NMR-Spektren zu den einzelnen Substrukturen durch formale Addition und Subtraktion auf den zugehörigen Strukturcodes erfolgen. Bei einer automatisierten Anwendung des Verfahrens unter Zuhilfenahme eines Computers können die entsprechenden Daten auf diese Weise besonders schnell verarbeitet werden, insbesondere, wenn die unten näher beschriebenen Cluster verwendet werden.

- 5 Um den Informationsgehalt der NMR-Spektren auf die wesentlichen relevanten Merkmale zu reduzieren, ist bei einer weiteren bevorzugten Verfahrensvariante vorgesehen, daß von den mehrdimensionalen NMR-Spektren Peaklisten zur Definition der zugeordneten Bereiche von Signalen erstellt werden.

Ein besonders einfaches Standardverfahren zur Peakerkennung besteht darin, daß ein Datenpunkt des mehrdimensionalen, vorzugsweise zweidimensionalen NMR-Spektrums als "Peak" erkannt wird, wenn sein Wert größer als der von n Nachbardatenpunkten ist, wobei vorzugsweise $n = 8$.

- 10 In einer bevorzugten Weiterbildung werden benachbarte Peaks zu Clustern zusammengefaßt und mittels Clusteranalyse ausgewertet, wobei ein oder mehrere Cluster als Bereiche von charakteristischen Signalen den interessierenden Substrukturen zugeordnet werden. Damit wird eine flächenartige Definition der zugeordneten Signalbereiche ermöglicht. Diese Methode ist unempfindlich gegen die Feinstruktur der einzelnen Peaks, die dabei außer Acht gelassen werden kann. Die Clusteranalyse als solche ist beispielsweise beschrieben in K.-P. Neidig et al., Journal of Magnetic Resonance 89, Seiten 543 bis 552 (1990).

Eine besonders bevorzugte Weiterbildung dieser Verfahrensvariante zeichnet sich dadurch aus, daß jedem Cluster ein Bereich von charakteristischen Signalen im zweidimensionalen NMR-Spektrum zugeordnet wird, und daß eine Substruktur als erkannt gilt, wenn für alle ihr zugeordneten Bereiche gilt, daß das über den zugeordneten Bereich integrierte NMR-Signal einen vorgegebenen Schwellwert überschreitet. Damit ist eine sehr zuverlässige Bilderkennung, insbesondere ein Erfassen von Substrukturen im NMR-Spektrum möglich.

- 20 Der Schwellenwert kann im einfachsten Fall als Konstante gewählt werden. Vorteilhafterweise wird der Schwellwert proportional zum Integral des NMR-Signals über Bereiche gewählt, die anderen interessierenden Substrukturen zugeordnet sind. Dies ermöglicht eine Normierung des jeweiligen zugeordneten Signalbereiches, die insbesondere für eine Core-Substruktur wünschenswert ist. Der Schwellwert ist dann unabhängig von absoluten Meßwerten der Spektren.

Bei einer weiteren bevorzugten Verfahrensvariante schließlich ist vorgesehen, daß eine Tabelle erstellt wird, in der für die gemessenen NMR-Spektren mindestens der restlichen Kombinationen (Zeilen) für alle interessierenden Substrukturen (Spalten) vermerkt wird, ob sie erkannt wurde (+) oder nicht (-), wobei in drei weiteren Spalten die Summe der erkannten Substrukturen, eine Gesamtbewertung ("wahr" ggf. "falsch", ggf. "unklar") und der geforderte Gesamtcode angegeben wird. Damit wird das Gesamtergebnis der kombinatorischen Meßserie auf einen Blick übersehbar.

- 30 Weitere Vorteile der Erfindung ergeben sich aus der Beschreibung und der Zeichnung. Ebenso können die vorstehend genannten und die noch weiter aufgeführten Merkmale erfindungsgemäß jeweils einzeln für sich oder zu mehreren in beliebigen Kombinationen Verwendung finden. Die gezeigten und beschriebenen Ausführungsformen sind nicht als abschließende Aufzählung zu verstehen, sondern haben vielmehr beispielhaften Charakter für die Schilderung der Erfindung.

Die Erfindung ist in der Zeichnung dargestellt und wird anhand eines Ausführungsbeispiels näher erläutert. Es zeigen:

Fig. 1a eine schematische Darstellung einer aus den Substrukturen $A_x + B_y + C_z$ aufgebauten Kombination eines organischen Moleküls, wobei die Substrukturen ein gemeinsames Core bilden;

- 40 Fig. 1b eine schematische Darstellung einer aus den Substrukturen $A_x + B_y + C_z$ sowie einem Core-Molekül aufgebauten Kombination eines organischen Moleküls;

Fig. 2 die in den Spektren der Fig. 3 bis 7 verwendeten Substrukturen in der Darstellungsweise der organischen Chemie; und

die Fig. 3 bis 7 jeweils ein NMR-Spektrum einer Kombination von drei Substrukturen A_x, B_y, C_z mit einem Core-Molekül, nämlich

- 45 Fig. 3 $A_2 + B_2 + C_2$;

Fig. 4 $A_2 + B_1 + C_1$;

Fig. 5 $A_2 + B_1 + C_3$;

Fig. 6 $A_2 + B_2 + C_3$;

- 50 Fig. 7 $A_1 + B_2 + C_2$.

Chemische Verbindungen einer kombinatorischen Serie eignen sich besonders gut für eine automatische oder zumindest teilweise automatische Interpretation, da die Signale der Strukturen von kombinatorischen Serien formal zerlegt werden können in ein Core-Modul, das für alle Mitglieder der Serie gleich ist, und in einige wenige variable Moduln, die systematisch mit einer begrenzten Anzahl von strukturellen Fragmenten innerhalb der kombinatorischen Serie durchvariiert werden.

- 55 In Fig. 1a ist sehr schematisch ein solches organische Molekül dargestellt, das aus drei Substrukturen A_x, B_y und C_z aufgebaut ist, welche im Bereich ihrer gegenseitigen Bindungen einen in Fig. 1a gestrichelt dargestellten gemeinsamen Core-Abschnitt bilden.

Alternativ kann, wie in Fig. 1b dargestellt, das organische Gesamtmolekül aber auch aus einem eigenen Core-Untermolekül mit drei daran anhängenden Substrukturen A_x, B_y und C_z aufgebaut sein.

- 60 Diese Moleküle werden üblicherweise mit Hilfe von Strukturcodes beschrieben, die aus Substrukturelementen mit den entsprechenden Substrukturcodes A_x, B_y und C_z usw. bestehen. Die Indizes x, y, z bedeuten jeweils eine fortlaufende Nummerierung in natürlichen Zahlen (1, 2, 3, ...).

Derartige Substrukturelemente können, wie unten gezeigt wird, als Signale oder zugeordnete Signalbereiche in zweidimensionalen HSQC-Spektren identifiziert werden. Die für die in den Fig. 3 bis 7 gezeigten Beispiele verwendeten Substrukturen sind chemische Substanzen, die in Fig. 2 entsprechend der Darstellungsweise der organischen Chemie abgebildet sind. Dabei bedeuten:

(1) 4-Nitro-phenyl = B2

- (2) Phenyl = B1
- (3) 3,4-methylenedioxy-phenoxy = A2
- (4) 3-Hydroxy,4-bromo-phenoxy = A1
- (5) tert-Butyloxycarbonyl-piperazyl = C2
- (6) Morpholinyl = C1
- (7) 2-Methoxy-piperazyl = C3

5

Bei den NMR-Experimenten, deren Ergebnisse in den Fig. 3 bis 7 gezeigt sind, handelt es sich um zweidimensionale, sogenannte HSQC (hetero nuclear single quantum coherence) Experimente. Angewandt auf Protonen und Kohlenstoff zeigen die Signale in diesen Spektren die Korrelationen zwischen Kohlenstoffatomen und den an ihnen chemisch gebundenen Protonen, und zwar die Kohlenstoffsignale in Richtung δ_1 und die Protonensignale in Richtung δ_2 .

10

Das Auffinden der zugeordneten Bereiche von charakteristischen Signalen, die zu einer bestimmten Substruktur gehören, kann manuell oder automatisiert durchgeführt werden. In der Automation führt man formal algebraische Additionen und Subtraktionen auf den Strukturcodes durch, ordnet jedem Code den Wert 1 zu und behält durch anschließende Schwellwertbildung den gewünschten Substrukturcode übrig, etwa:

15

"A₂ B₂ C₂"
 "A₂ B₁ C₁"
 "A₂ B₁ C₃"
 "A₂ B₂ C₃"
 "A₁ B₂ C₂"

20

dann kann durch Addition von

"A₂ B₁ C₃"
 "A₂ B₂ C₃" und

25

Subtraktion

"A₂ B₂ C₂"
 "A₂ B₁ C₁"
 "A₁ B₂ C₂"

30

das Ergebnis für A₁ A₂ B₁ B₂ C₁ C₂ C₃ -1 0 0 -1 -1 -2 2 erhalten werden. Setzt man einen Schwellwert 2, dann bleibt nur noch C₃ übrig.

Die allgemeine Regel lautet: Addiere alle Strukturcodes, die den gewünschten Substrukturcode enthalten, z. B. N, dann subtrahiere alle anderen, dann setze den Schwellwert auf N.

35

Diese formale Operation wird in der Praxis auf Peaklisten, die von den entsprechenden Spektren gewonnen wurden, durchgeführt. Ein Datenpunkt wird als zweidimensionaler Peak erkannt, wenn sein Wert größer als jeder der seiner 8 nächsten Nachbarn ist. Benachbarte Peaks können mit Hilfe einer Clusteranalyse, die paarweise Distanzen und Intensitäten auswertet, zu Gruppen (Cluster) zusammengefaßt werden.

40

Wird ein Peak eines Spektrums zu den Peaks eines anderen Spektrums addiert, so wird er in die dortige Liste aufgenommen und erhöht die Intensitäten aller Peaks, die innerhalb eines definierten Radius liegen.

Wird ein Peak eines Spektrums von den Peaks eines anderen Spektrums subtrahiert, so wird er aus der dortigen Liste entfernt und erniedrigt die Intensitäten aller Peaks, die innerhalb eines definierten Radius liegen.

Als Ergebnis erhält man eine Liste von Peaks, die von den Signalen der gewünschten Substruktur herrühren. Da diese Signale in verschiedenen Spektren leicht unterschiedlich sein können, treten die Peaks mehrfach und in Gruppen auf. Die Gruppen, auch Cluster genannt, werden durch eine Clusteranalyse bestimmt.

45

Die für die gewünschte Substruktur erhaltenen Signale werden nun durch jeweils kleine Rechtecke repräsentiert, wobei jedes Rechteck genau ein Cluster enthält. Die Breite und Höhe dieser Areale entsprechen den erwarteten Variationen der Signale im gegebenen Satz von Spektren.

50

Bei der Überprüfung der restlichen Spektren wird eine Integration aller Substrukturen durchgeführt. Integration bedeutet hierbei Summation aller entsprechenden Datenpunkte. Außerdem wird in jedem Spektrum ein vordefiniertes Gebiet, welches keine Signale enthält, integriert und daraus ein Rauschwert berechnet. Der Rauschwert wird von allen Integralen abgezogen.

Optional können Signale der "Core"-Substruktur als Referenz definiert und gesondert integriert werden.

55

Optional können die Integralverhältnisse zwischen allen Arealen aller Substrukturen berechnet werden.

Eine Substruktur gilt als erkannt, wenn alle ihre Areale den Integrationswert > 0 annehmen.

Optional gilt eine Substruktur als erkannt, wenn alle ihre Areale im Verhältnis zu einer Referenz einen definierten Integrationswert überschreiten.

Optional gilt eine Substruktur als erkannt, wenn alle Verhältnisse der Integrale aller ihrer Areale zu den Integralen aller anderen Areale aller anderen Substrukturen einen definierten Wert überschreiten.

60

Im gezeigten Ausführungsbeispiel war x = 2, y = 2 und z = 3. Daraus ergeben sich 2 · 2 · 3 = 12 kombinatorisch mögliche Kombinationen (A_x B_y C_z). Der minimale repräsentative Satz müßte mindestens drei Kombinationen umfassen, damit C₁, C₂ und C₃ jeweils repräsentiert sind.

Als Resultat der Messungen und ihrer Auswertung ergibt sich bei unserem Ausführungsbeispiel folgende Ergebnistabelle:

65

Kombinationen	A ₂ B ₂ C ₂	A ₂ B ₁ C ₁	A ₂ B ₁ C ₃	A ₂ B ₂ C ₃	A ₁ B ₂ C ₂
A ₁	-	-	-	-	+
A ₂	+	+	+	+	-
B ₁	-	+	+	-	-
B ₂	+	-	-	+	+
C ₁	-	+	-	-	-
C ₂	+	-	-	-	+
C ₃	-	-	+	+	-
Core	+	+	+	+	+
Summe	4	4	4	4	4
Resultat	+	+	+	+	+

Patentansprüche

1. Verfahren zum Verifizieren der Synthese von aus Substrukturen aufgebauten organischen Molekülen der kombinatorischen Chemie, wobei die Substrukturen zumindest Teilbereiche der Ausgangssubstanzen enthalten, mit folgende Schritten:

(A) Festlegung der Substrukturen der organischen Moleküle aufgrund der Ausgangssubstanzen und Zuordnen von Substrukturcodes (A_x, B_y, C_z, ...);

(B) Zuordnung eines Gesamtcodes für die jeweiligen Kombinationen von aus unterschiedlichen Substrukturen aufgebauten synthetisierten organischen Molekülen;

(C) Aufnahme von mehrdimensionalen NMR-Spektren der durch die jeweilige Kombination erhaltenen synthetisierten organischen Moleküle für zumindest eine Teilmenge der kombinatorisch möglichen Moleküle;

(D) eindeutige Zuordnung von Bereichen im NMR-Spektrum, in denen charakteristische Signale einer Substruktur vorkommen zu dieser Substruktur;

(E) Überprüfung der NMR-Spektren der Kombinationen auf das Vorhandensein der charakteristischen Signale in allen zugeordneten Bereichen;

(F) falls die Überprüfung für eine bestimmte Kombination ergibt, daß genau in den zugeordneten Bereichen die Signale der in ihrem Gesamtcode enthaltenen Substrukturen im NMR-Spektrum beobachtet wurden, wird diese Kombination mit "wahr" charakterisiert.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß sich der Schritt (G) anschließt:

(G) falls die Überprüfung für eine bestimmte Kombination ergibt, daß in mindestens einem zugeordneten Bereich die charakteristischen Signale der in ihrem Gesamtcode enthaltenen Substrukturen im NMR-Spektrum nicht beobachtet wurden, wird diese Kombination mit "falsch" charakterisiert.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß sich der Schritt (H) anschließt:

(H) falls die Überprüfung für eine bestimmte Kombination ergibt, daß weder (F) noch (G) zutrifft, wird diese Kombination mit "unklar" charakterisiert.

4. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Schritte (D) und (E) durch folgende Schritte realisiert werden:

(D₁') Auswahl eines repräsentativen Satzes von Kombinationen, der alle interessierenden vorkommenden Substrukturen zumindest einmal enthält;

(D₂') eindeutige Zuordnung von Bereichen mit charakteristischen Signalen der NMR-Spektren des repräsentativen Satzes zu den einzelnen Substrukturen;

(D₃') falls sich nicht jeder Substruktur eindeutig ein Bereich von Signalen zuordnen läßt, erfolgt eine Auswahl eines geänderten Satzes von Kombinationen und Wiederholung ab Schritt (D₂');

(E') Überprüfung der NMR-Spektren der restlichen, nicht dem repräsentativen Satz angehörigen Kombinationen auf das Vorhandensein aller zugeordneten Bereiche von Signalen.

5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß der repräsentative Satz in Schritt (D₁') durch Modifikation eines repräsentativen Satzes von Kombinationen aus einer früheren Meßreihe abgeleitet wird.

6. Verfahren nach Anspruch 4 oder 5, dadurch gekennzeichnet, daß die Anzahl der im repräsentativen Satz enthaltenen Kombinationen minimiert wird.

7. Verfahren nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet, daß die Zuordnung von Bereichen von Si-

gnalen der NMR-Spektren des repräsentativen Satzes zu den einzelnen Substrukturen in Schritt (D₂') manuell erfolgt.

8. Verfahren nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet, daß die Zuordnung von Bereichen von Signalen der NMR-Spektren des repräsentativen Satzes zu den einzelnen Substrukturen in Schritt (D₂') automatisiert mittels eines Rechners erfolgt.

9. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß die Zuordnung von Bereichen von Signalen zu den Substrukturen in Schritt (D) aufgrund von NMR-Spektren aus einer früheren Meßreihe und/oder aufgrund theoretisch berechneter Spektraldaten erfolgt.

10. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die organischen Moleküle niedermolekular sind, vorzugsweise mit Molekulargewichten im Bereich von 100 u bis 2000 u.

11. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß ein Teil des organischen Gesamtmoleküls als Core in allen Kombinationen identisch vorhanden ist.

12. Verfahren nach Anspruch 11, dadurch gekennzeichnet, daß das Core eine Substruktur ist, die zwischen zwei und sechs chemische Bindungsstellen aufweist.

13. Verfahren nach Anspruch 11 oder 12, dadurch gekennzeichnet, daß die Substrukturen zu Gruppen zusammenfaßbar sind, die jeweils einer Bindungsstelle des Core zugeordnet sind.

14. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Anzahl der interessierenden Substrukturen zwischen fünf und fünfhundert liegt.

15. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß das mehrdimensionale NMR-Spektrum ein zweidimensionales, vorzugsweise ein HSQC-Spektrum ist.

16. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß das mehrdimensionale NMR-Spektrum die Atomkerne ¹H und ¹³C umfaßt.

17. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Zuordnung von Bereichen von Signalen der NMR-Spektren zu den einzelnen Substrukturen durch formale Addition und Subtraktion auf den zugehörigen Strukturcodes erfolgt.

18. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß von den mehrdimensionalen NMR-Spektren Peaklisten zur Definition der zugeordneten Bereiche von Signalen erstellt werden.

19. Verfahren nach Anspruch 18, dadurch gekennzeichnet, daß ein Datenpunkt des mehrdimensionalen, vorzugsweise zweidimensionalen NMR-Spektrums als Peak erkannt wird, wenn sein Wert größer als der von n Nachbardatenpunkten ist, wobei vorzugsweise n = 8.

20. Verfahren nach einem der Ansprüche 18 oder 19, dadurch gekennzeichnet, daß benachbarte Peaks zu Clustern zusammengefaßt und mittels Clusteranalyse ausgewertet werden, wobei ein oder mehrere Cluster als Bereiche von Signalen den interessierenden Substrukturen zugeordnet werden.

21. Verfahren nach Anspruch 20, dadurch gekennzeichnet, daß jedem Cluster ein Bereich von charakteristischen Signalen im zweidimensionalen NMR-Spektrum zugeordnet wird, und daß eine Substruktur als erkannt gilt, wenn für alle ihr zugeordneten Bereiche gilt, daß das über den zugeordneten Bereich integrierte NMR-Signal einen vorgegebenen Schwellwert überschreitet.

22. Verfahren nach Anspruch 21, dadurch gekennzeichnet, daß der Schwellwert proportional ist zum Integral des NMR-Signals über Bereiche, die anderen interessierenden Substrukturen zugeordnet sind.

23. Verfahren nach einem der vorangehenden Ansprüche, dadurch gekennzeichnet, daß eine Tabelle erstellt wird, in der für die gemessenen NMR-Spektren mindestens der restlichen Kombinationen in Zeilen für alle interessierenden Substrukturen in Spalten vermerkt wird, ob sie erkannt wurde (+) oder nicht (-), wobei in drei weiteren Spalten die Summe erkannten Substrukturen, eine Gesamtbewertung als "wahr", ggf. "falsch", ggf. "- unklar" und der geforderte Gesamtcode angegeben wird.

Hierzu 6 Seite(n) Zeichnungen

- Leerseite -

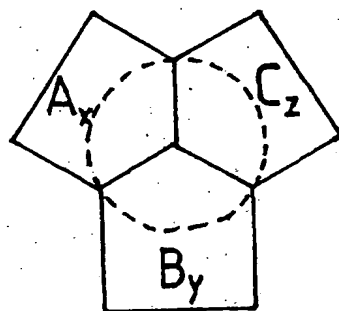


Fig. 1a

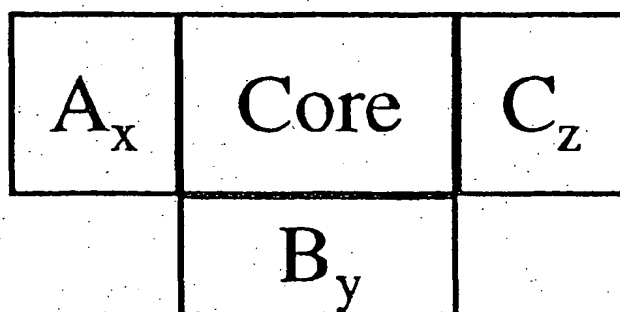


Fig. 1b

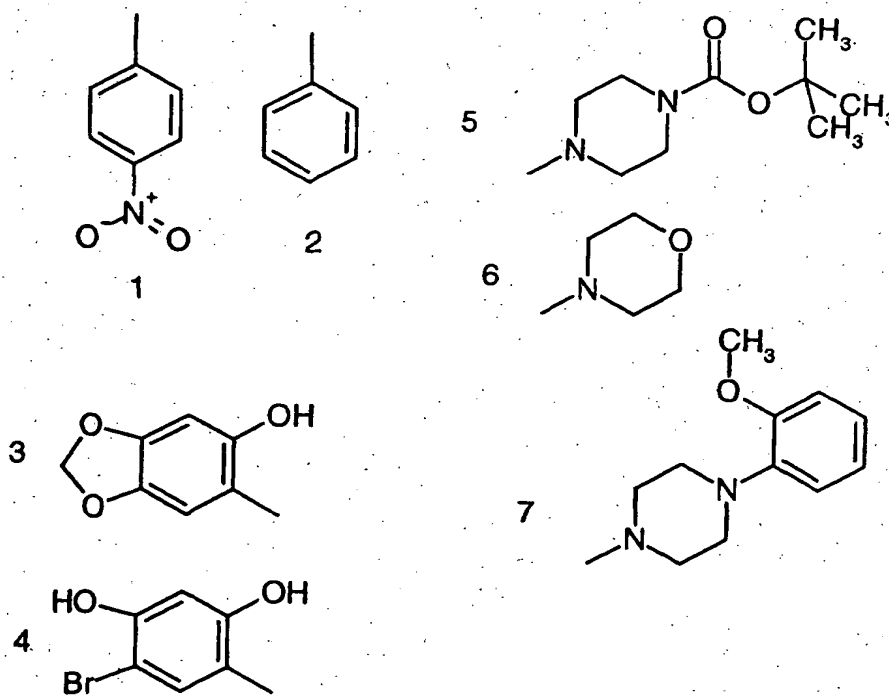


Fig. 2

$A_2 B_2 C_2$

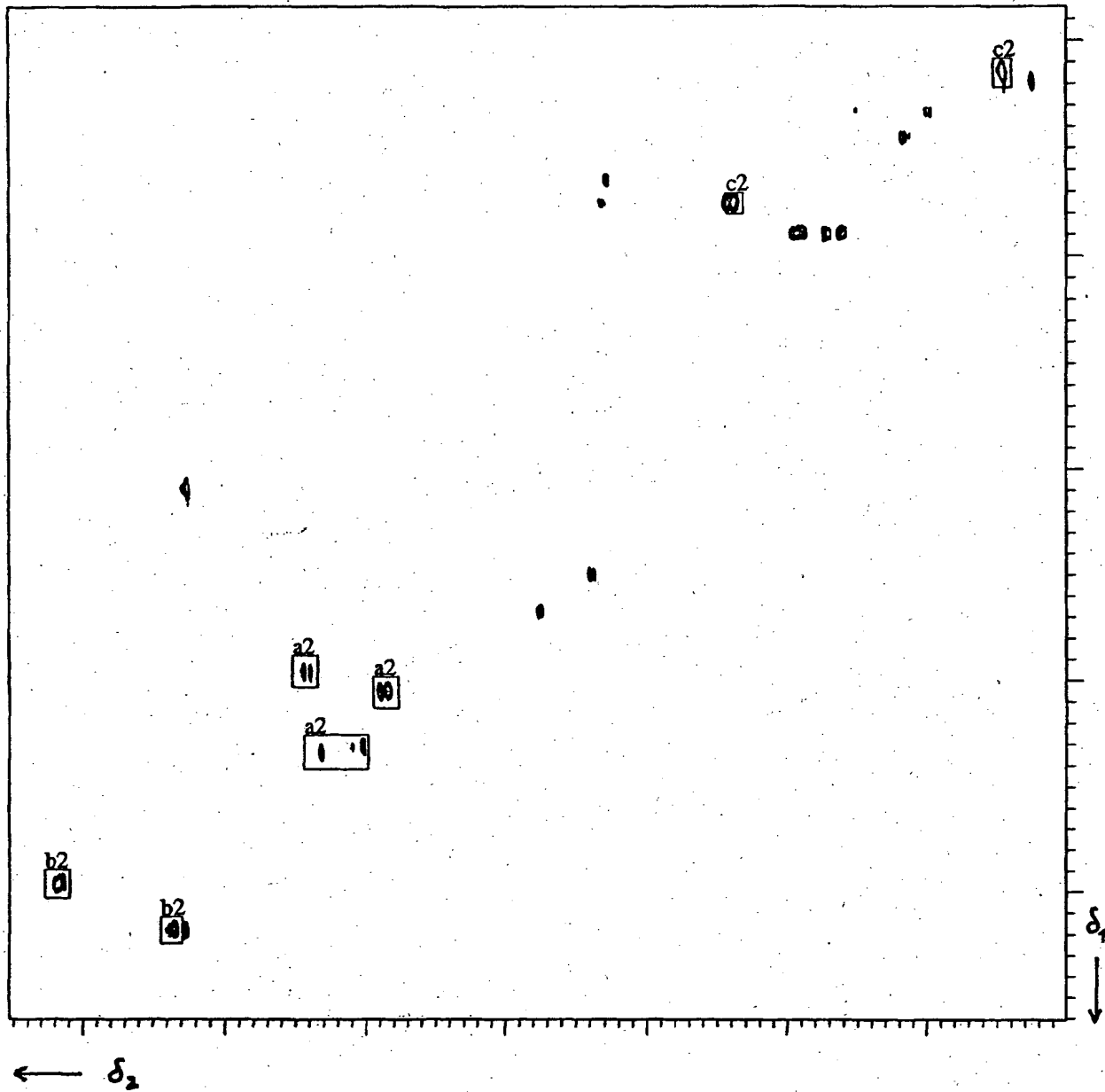


Fig. 3

$A_2 B_1 C_1$

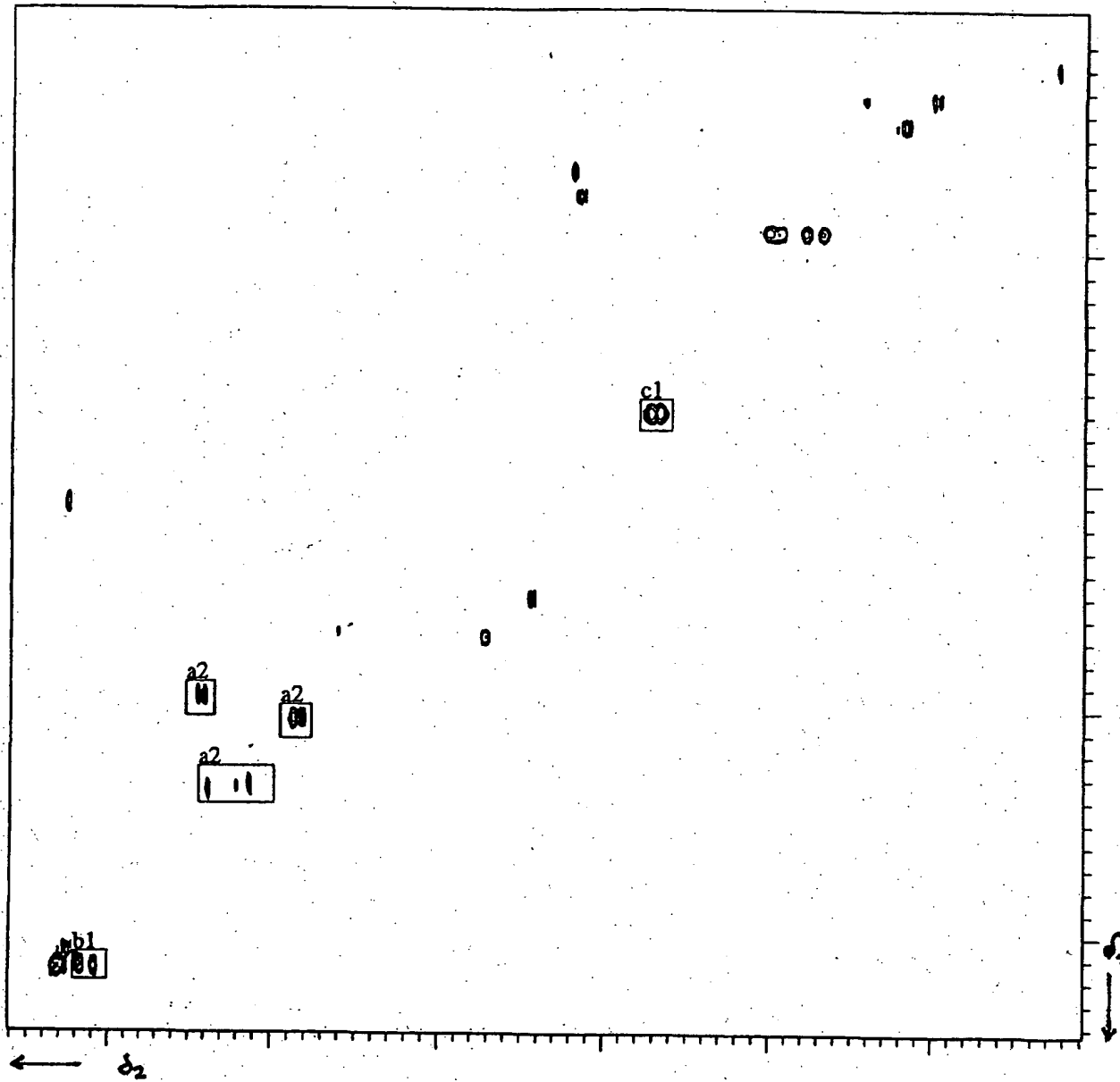


Fig. 4

A_2 B_1 C_3

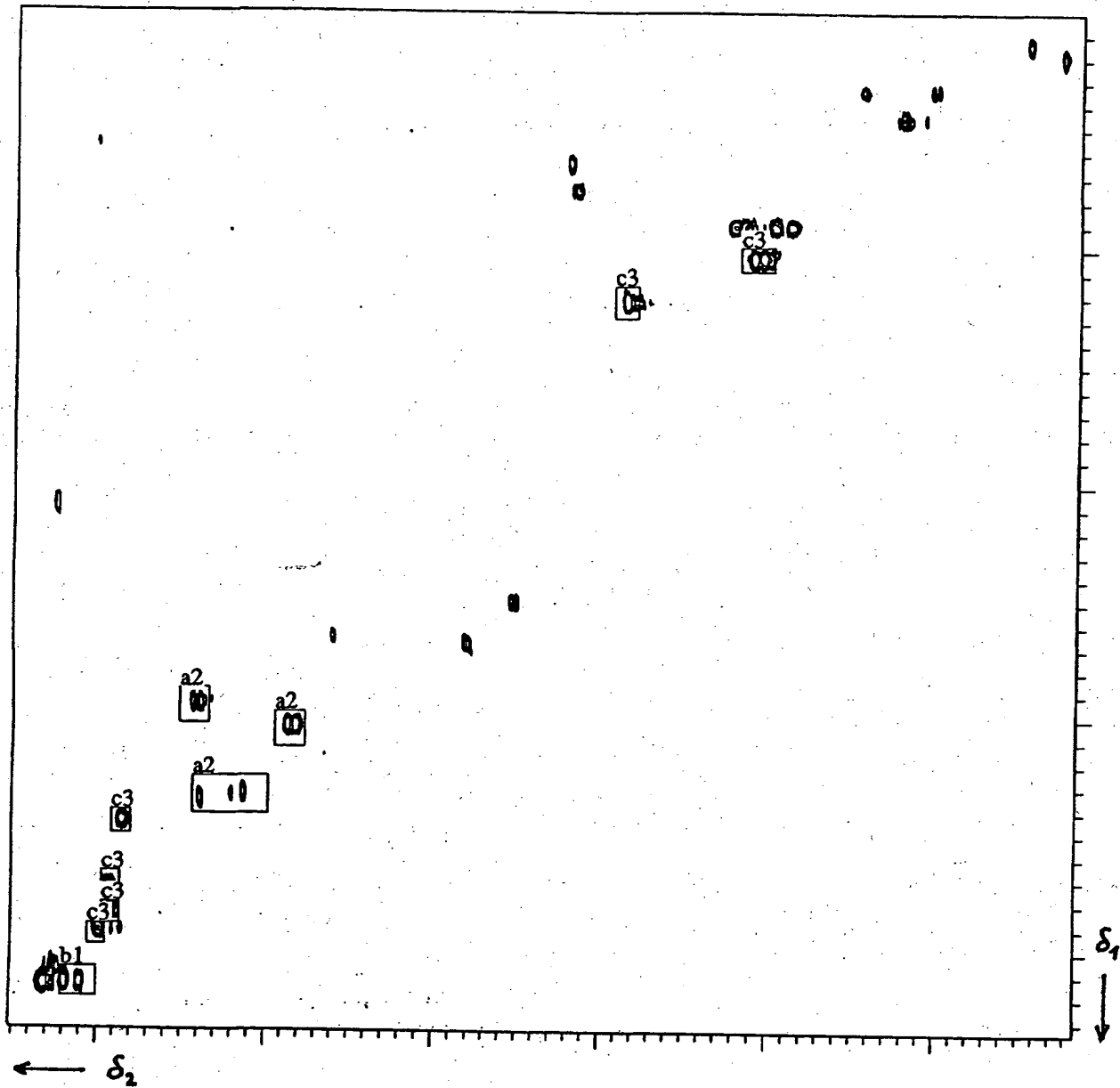


Fig. 5

$A_2 B_2 C_3$

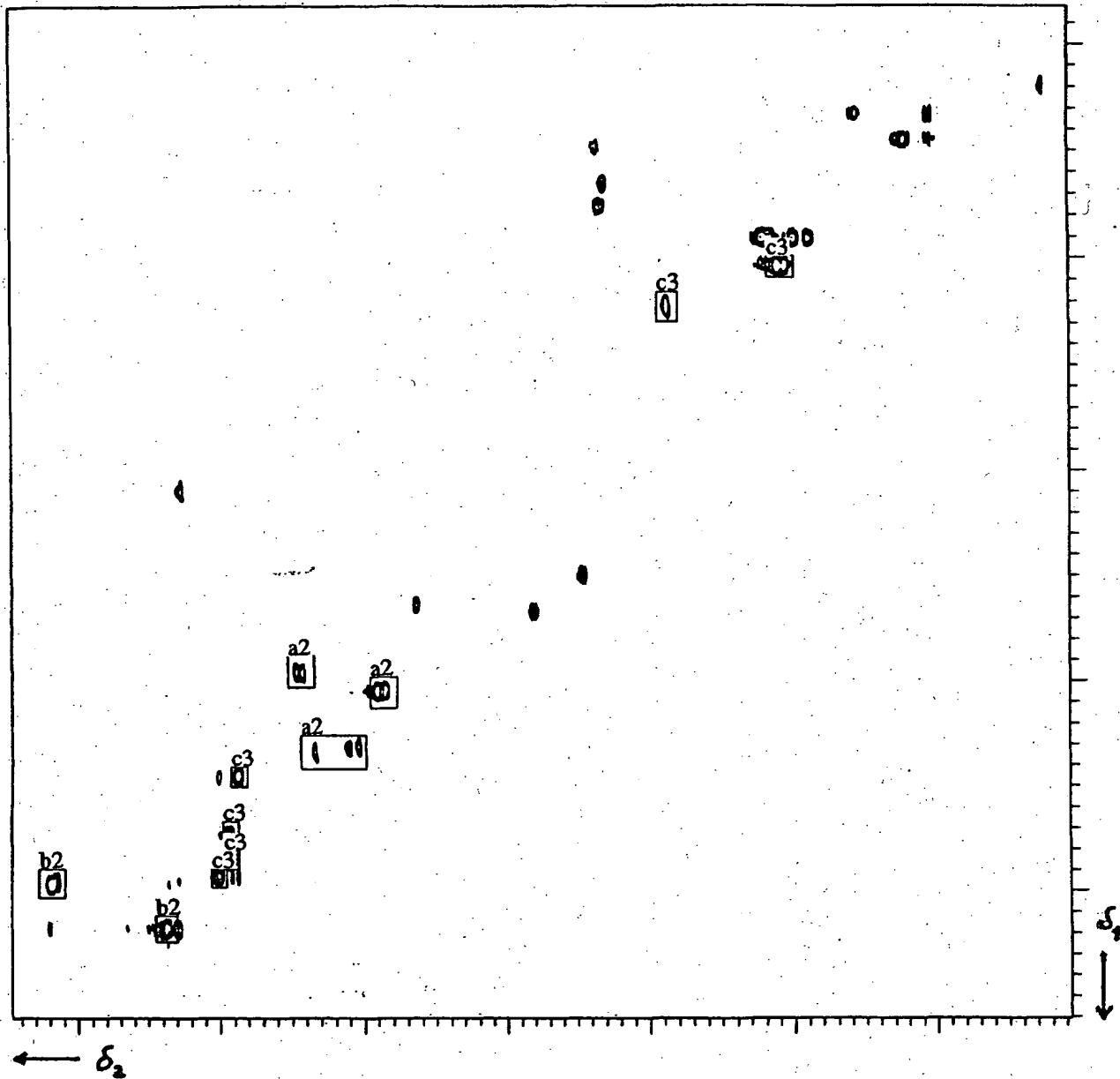


Fig. 6

$A_1 B_2 C_2$

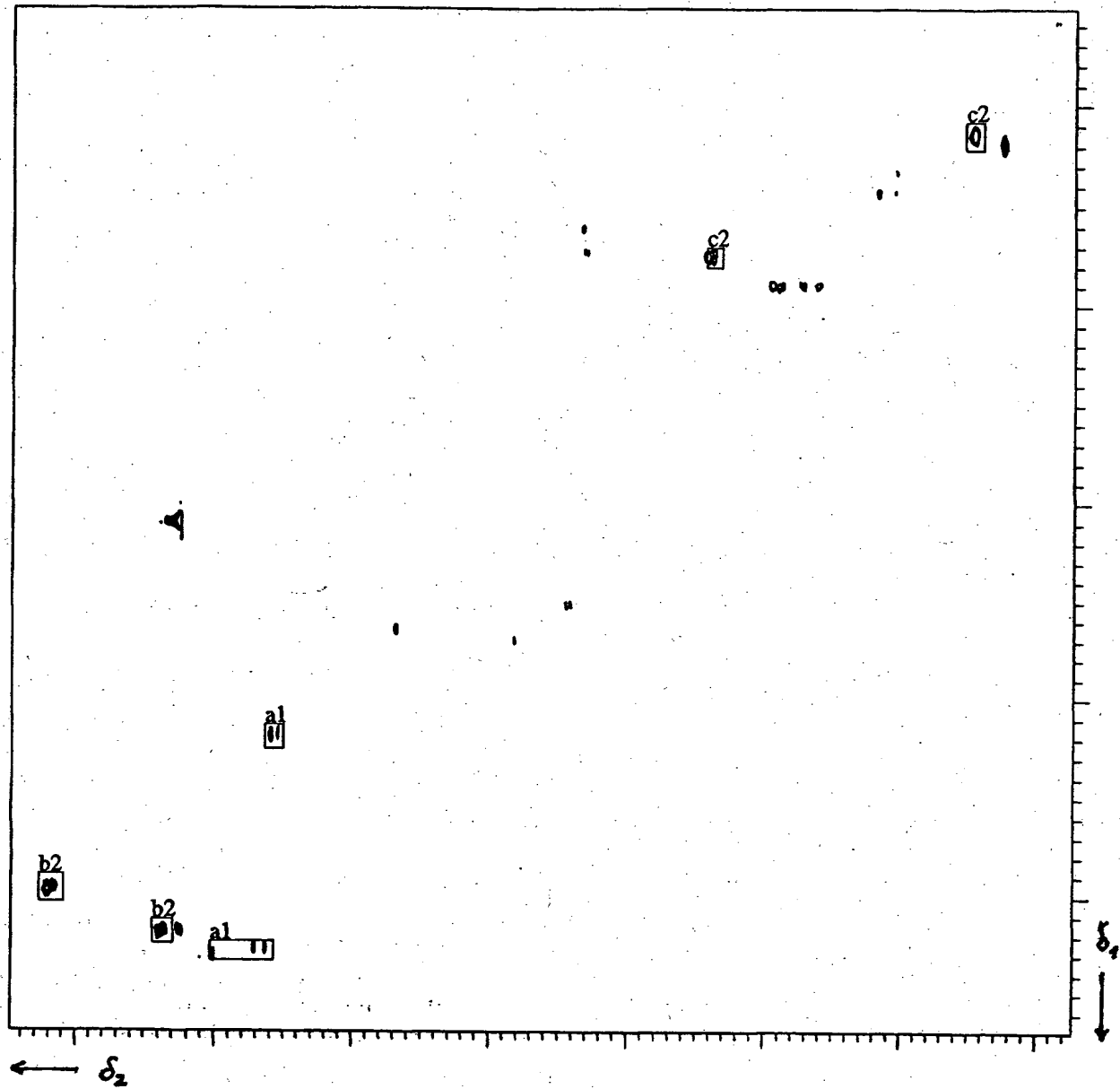


Fig. 7